Radial Basis Function Neural Network for Work Zone Capacity and Queue Estimation

Asim Karim[1] and Hojjat Adeli, F.ASCE[2]

Abstract: An adaptive computational model is presented for estimating the work zone capacity and queue length and delay, taking into account the following factors: number of lanes, number of open lanes, work zone layout, length, lane width, percentage trucks, grade, speed, work intensity, darkness factor, and proximity of ramps. The model integrates judiciously the mathematical rigor of traffic flow theory with the adaptability of neural network analysis. A radial-basis function neural network model is developed to learn the mapping from quantifiable and nonquantifiable factors describing the work zone traffic control problem to the associated work zone capacity. This model exhibits good generalization properties from a small set of training data, a specially attractive feature for estimating the work zone capacity where only limited data is available. Queue delays and lengths are computed using a deterministic traffic flow model based on the estimated work zone capacity. The result of this research is being used to develop an intelligent decision support system to help work zone engineers perform scenario analysis and create traffic management plans consistently, reliably, and efficiently.

DOI: 10.1061/(ASCE)0733-947X(2003)129:5(494)

CE Database subject headings: Neural networks; Models; Workspace; Highway construction; Traffic management.

Introduction

Recognizing the need for serving the public's present and future transportation needs, the Transportation Equity Act for the 21st Century (FHWA 1998) has earmarked increased funding for maintenance, rehabilitation, and reconstruction of the nation's aging highway system. It is therefore expected that the number of work zones would increase in the future, further impacting the normal operation of the highway system. The primary goal of the various departments of transportation (DOTs or traffic agencies) is to enhance mobility and safety at all times on the highway system. Over the years, the quality and uniformity of traffic control devices and procedures have improved. However, the numbers of work zone related fatalities and injuries have remained practically the same, and the traveling public has become increasingly frustrated with additional mobility restrictions (FHWA 2000). A large population is exposed to work zones and their negative impacts. This in turn generates widespread negative sentiments towards the public agencies responsible for providing efficient and safe transportation services to the public.

As work zones on today's highways are becoming an increasingly frequent and unavoidable reality, traffic agencies are faced with the challenging problem of effectively planning and managing work zones in their jurisdictions. This is a complex, multifaceted problem that requires life-cycle cost analyses at the system level. An analysis at the system level, however, will only be useful when a reliable model for the impact of a given work zone on traffic flow is available. In current practice, work zone engineers rely on their judgment based on previous experiences to quantify work zone traffic impacts and to make decisions. Work zone engineers have to consider a large number of factors such as work zone layout, work intensity, diversion of traffic, and driver behavior.

The effectiveness of a work zone traffic management plan (TMP) may be measured by the delay experienced by motorists and/or the length of queue formed on the upstream side. To improve the objectivity and reliability of a work zone TMP, a reliable model is needed that maps traffic flow and work zone characteristics to delay time and queue length. For such a model to be useful in practice, it must have the following characteristics:

- It should be based on a simple underlying principle of traffic flow. Complicated physical and/or psychological models of traffic flow are unrealizable and intractable for practical purposes. Also, the data input needed for some of these models are not readily available, thus introducing a source of error. A simple model, on the other hand, can be reasoned with and "calibrated" to produce reliable results for different work zone scenarios.
- It should consider the major factors that affect traffic flow through work zones. For example, work zone capacity should not be an input, but rather should be determined from an input of work zone characteristics. Consequently, the model should be able to process both quantifiable and nonquantifiable (or linguistic) variables involved in the analysis.
- It should be flexible in the sense that it can be adapted and extended for the analysis of different work zone traffic control scenarios. In particular, its applicability should not be restricted to a single roadway geometry and/or work zone layout. To accomplish this, the model should be capable of learning from input/output data and not be based solely on a

[1]Assistant Professor, Computer Science Dept., Lahore Univ. of Managment Sciences, Pakistan.

[2]Professor, Dept. of Civil and Environmental Engineering and Geodetic Science, The Ohio State Univ., 470 Hitchcock Hall, 2070 Neil Ave., Columbus, OH, 43210.

Note. Discussion open until February 1, 2004. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on November 20, 2001; approved on October 21, 2002. This paper is part of the ***Journal of Transportation Engineering***, Vol. 129, No. 5, September 1, 2003. ©ASCE, ISSN 0733-947X/2003/5-494–503/$18.00.

physical/psychological model of traffic flow. This is essential because real world behavior under all situations can not be modeled satisfactorily by conventional mathematics only.

With these guidelines in mind, a new adaptive computational model is developed for estimating the work zone capacity as well as queue delay and length based on the conservation principle of traffic flow. The model integrates judiciously the mathematical rigor of traffic flow theory with the adaptability of neural network analysis. A radial-basis function neural network (RBFNN) model is developed to learn the mapping from quantifiable and nonquantifiable factors describing the work zone traffic control problem to the associated work zone capacity. Queue delays and lengths are computed using a deterministic traffic flow model based on the estimated work zone capacity. The goal of this research is to help create an intelligent decision support system to help work zone traffic engineers create TMPs consistently and efficiently.

Review of Work Zone Traffic Control Planning

A work zone is a region within an existing highway's roadway where active maintenance, rehabilitation, and/or reconstruction work is carried out. The highway is not closed, and traffic flow and highway work exist in close proximity to each other. A work zone thus imposes a spatial and temporal restriction on a highway's roadway that negatively impacts the normal flow of traffic. These impacts appear as increased congestion, travel times, accidents, and a greater level of dissatisfaction among the traveling public. Work zones are planned and managed to minimize these impacts and the overall cost. A primary concern of traffic agencies is creation of TMPs for long-term stationary work zones (with duration of more than one day), because they are of high impact and visibility with a lot at stake for all parties involved. Development of such plans requires a careful analysis of traffic flow through the work zone to determine the best work phasing and work zone layout. As mentioned earlier, the overall problem of planning and managing of work zones in a highway system is complex, requiring life-cycle cost analyses. Nowadays, however, the primary focus of traffic agencies is the creation of a TMP for a given work zone that minimizes queue delays and lengths.

Construction and maintenance work zones on highways have been studied for more than 30 years in an effort to develop safer and more effective TMPs. A survey of the literature that specifically deals with mobility of traffic through work zones reveals a mix of empirical studies and mathematical analyses. Empirical studies collect and analyze data from work zones in an effort to develop an understanding of traffic demand, work zone capacity, work zone layout, traffic mitigation strategies, and traffic congestion.

By analyzing data from 161 observations of freeway queuing, Cottrell (2001) presents an empirical model of queuing delay using linear regression analyses. Equations are presented that relate traffic flow and capacity variables to queue delay variables. The model, however, is for recurrent congestion only and does not consider congestion caused by work zones and their associated variables. Cassidy and Mauch (2001) also study recurrent congestion using cumulative plots of traffic count and show that the density in long queues that span several interchanges decreases in the upstream direction. In an earlier study, Cassidy and Bertini (1999) analyze discharge patterns from freeway bottlenecks and conclude that discharge rates are nearly constant when taken cumulatively. These two articles provide general insights into queuing behavior caused by bottlenecks that may be applicable to work zones also. By analyzing data from 24 work zones, Dixon et al. (1996) present capacity values of work zones for both urban and rural freeways. They found that work zone capacity is affected significantly by work intensity, rural and urban location, and darkness. The presence of a work zone forces many regular motorists to choose alternative routes even when a diversion is not specified explicitly. This phenomenon, called natural diversion, can play a significant role in work zone traffic flow analysis. Natural diversion is studied by Ullman (1996) for the particular situation where the freeway has continuous frontage roads. Krammes and Lopez (1994) provide recommendations for work zone capacity after analyzing data from 33 work zones. They present a single base work zone capacity for different work zone configurations. This base value can be modified to reflect the effects of work intensity, traffic composition (percentage of trucks), and the presence of ramps just upstream of the work zone. The *Highway Capacity Manual* (TRB 2000) represents the current state of practice in traffic analyses, summarizing the results of empirical studies of the past 20 years. Its coverage of work zone capacity, however, is brief and limited to a few general recommendations.

Empirical studies are generally limited in scope and not readily applicable to decision making where different scenarios have to be analyzed objectively. They do provide valuable insights, but such insights are case-specific and have to be captured in a generalized computational model to be of value to traffic agencies in the development of TMPs. Several models have been proposed in the literature for the determination of queue lengths and delay times associated with work zones. In general, these models are based on one of two approaches of traffic flow theory: shock wave analysis and queuing analysis (May 1990). Shock wave analysis traces shock waves, or boundaries demarcating different flow regimes, in time and space to determine regions of queued (congested) and uncongested flow. Shock wave analysis is deterministic in nature and uses only two macroscopic traffic flow parameters, traffic demand and roadway capacity, as input. It does not take into account the dependence of traffic demand and work zone capacity on many other parameters and is therefore not very reliable for work zone traffic flow analysis. In queuing analysis, a work zone is modeled as a service process where vehicles arriving at the work zone experience some delay (queue delay) before they are able to continue along the highway. Queuing analysis can be deterministic or stochastic, and it may use either macroscopic or microscopic traffic flow parameters. Deterministic queuing analysis suffers from the same shortcoming as shock wave analysis, while stochastic analysis requires traffic distribution information that is often not available in practice. A microscopic stochastic model of traffic flow provides the most detailed analysis possible. The accuracy of such a model, however, depends on the accuracy of human-vehicle-environment behavior models, which are still not well understood and are an area of research.

Chien and Schonfeld (2001) present an optimization model for the optimal length of a work zone on a four-lane divided highway (two lanes in each direction) with one lane in each direction closed. The objective function is the total cost, including user cost, accident cost, and agency maintenance cost. The model assumes that work zone capacity is constant and independent of work zone characteristics. Islam and Seneviratne (1993) evaluate the suitability and effectiveness of four traffic planning software tools for the evaluation of TMPs, while Sadegh et al. (1988) present a simulation model for work zones on arterials. The computer program QUEWZ (queue and user cost evaluation of work zones) evaluates queue lengths and additional user costs for work zones on freeways (Memmott and Dudek 1984; Krammes et al. 1987). This model uses the conservation of flow principle to calculate the queue lengths and user costs for different lane closure configurations and work schedules. The capacity of the work zone is calculated from empirical speed-flow-density relationships and is independent of the work zone characteristics, such as work zone layout and work intensity.

Recently, an initiative at the Federal Highway Administration (FHWA) was launched to develop strategic tools for work zone traffic analysis and decision support. As part of this program, a spreadsheet based tool called QuickZone has been developed to quantify work zone queue delays and lengths given work zone capacity, traffic demand, and work phasing (Mitretek 2000). The QuickZone software takes as input hourly values of traffic demand and work zone capacity. Queue lengths and delays are computed by the deterministic input-output conservation principle of traffic flow. The software does not take into account work zone layout, lane widths, driver behavior, work intensity, and proximity of ramps in the computation of work zone queue delays and lengths.

Deterministic Queuing Model for Work Zone Traffic Analysis

A deterministic macroscopic queuing model is used to calculate queue lengths and delays produced by bottlenecks on highways such as work zones. This model is based on the principle of conservation of flow, which states that under homogeneous roadway conditions the number of vehicles entering a segment in a given time period must be equal to the number of vehicles exiting the segment in the same time period. If the road segment is inhomogeneous, with a bottleneck such as a capacity reducing work zone existing on a portion of the segment, then the number of vehicles exiting may be fewer than the number of vehicles entering the segment. The difference in such a situation represents the queue formed in the upstream direction.

This model is an adaptation of the theory of incompressible fluid flow to vehicular traffic streams. It only requires as input traffic demand (or flow rate), highway capacity, and their variation over time. If conservation of flow is evaluated over a reasonable time period (say, greater than 15 min) the model produces practically accurate estimates for queue lengths and delay times that can be used for planning, assuming that the demand and capacity values represent the actual conditions on the highway accurately. In the following paragraphs, the model is formulated for a single link or segment of freeway containing a capacity reducing work zone.

Fig. 1 shows the layout of a freeway segment with a construction work zone. The work zone acts like a metered on-ramp that allows only a certain number of vehicles to pass through in a given amount of time. Major freeway repair, rehabilitation, and reconstruction projects are multiday repetitive operations where work zone layout and phasing is often identical from day to day. Therefore, it is sufficient to analyze a typical day (or at most a few typical days) of work for their traffic impact. If a 1 h evaluation time period ($\Delta t = 1$ h) is considered, 24 values of traffic demand or anticipated hourly traffic flow approaching the work zone (f_t) and highway capacity (c) are needed as input to cover a period of 1 day. Let these values be denoted by f_t ($t=0,..,23$) and $c(t)$ ($t=0,..,23$), respectively, where the index t indicates the time period. Then, for time periods $t=1, 2, 3,..., 23$, using the conservation principle, the number of vehicles in the queue in time period t is given by

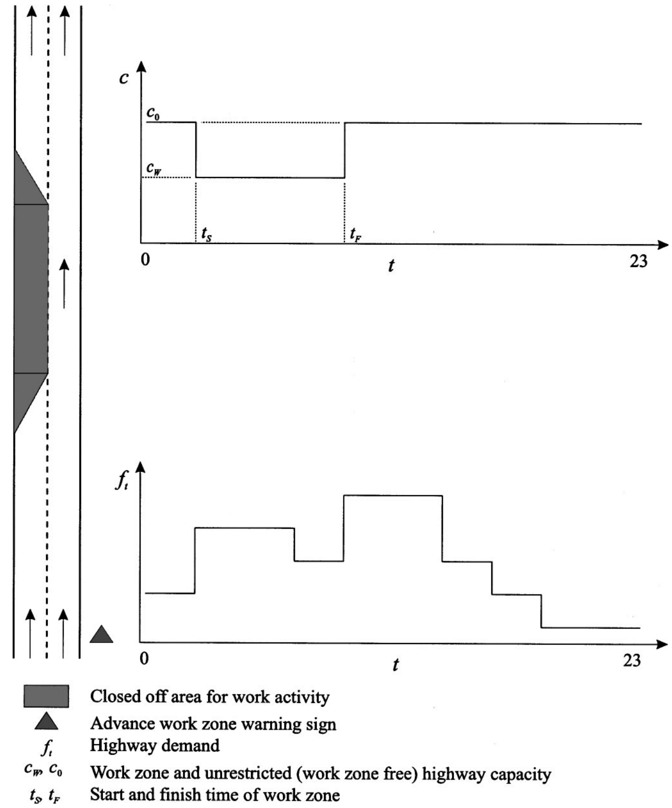

Fig. 1. Schematic description of traffic demand and capacity through work zones

$$Q(t) = \max[f_t(t) - c(t) + Q(t-1), 0] \quad t = 1,2,3,\ldots,23 \tag{1}$$

The term $Q(t-1)$ on the right-hand side of recursive Eq. (1) represents the number of vehicles in the queue for the previous time period. When $t=1$, the queued vehicles at the previous time period $Q(t-1)$ can be taken equal to zero if the beginning of the evaluation period is chosen at a time when demand is less than the capacity. The highway capacity $c(t)$ is equal to the work zone capacity, c_W, when the work zone exists and equal to c_0, the capacity of the freeway, in the absence of the work zone (Fig. 1).

If k is the jam density (the number of queued vehicles that occupy a given length of highway), then the average length of the queue at time period t is given by

$$Q_L(t) = \frac{Q(t)}{k} \quad t = 0,1,2,\ldots,23 \tag{2}$$

The expected daily queue delay (in vehicle-hours) experienced by motorists is given by

$$Q_D = \sum_{t=1}^{23} \frac{Q(t) + Q(t-1)}{2} \Delta t \tag{3}$$

where $\Delta t = 1$ h is the evaluation time period. Eqs. (1)–(3) define the primary parameters needed by the work zone engineer to assess the impact of a work zone on traffic in creating a TMP. Using these values, the user delay cost can be estimated provided that the average cost per vehicle hour expressed in dollars per vehicle hour is known. The accuracy of this analysis depends on the accuracy of the demand and capacity values used.

Estimation of Traffic Demand at Work Zones

To accurately predict the temporal development and extent of queues and delays created in work zones, it is necessary to have a reasonably accurate estimate of the traffic demand—that is, the anticipated hourly traffic flow of the freeway approaching the work zone. The annual average daily traffic (AADT), which is the daily traffic demand averaged over all days of the year, is unsuitable for this purpose. Shorter time estimates of traffic demand can be obtained from the AADT when daily, weekly, and seasonal demand factors (α_S) are known for the freeway. Many traffic agencies maintain values for daily, weekly, and seasonal demand factors. For example, the Ohio Department of Transportation (ODOT) specifies values for α_S in the range of 0.76 and 1.72 (http://www.dot.state.oh.us/techservsite/). Many traffic agencies also maintain hourly vehicle counts on major highways, in which case $\alpha_S = 1$.

These traffic demand values reflect the behavior and usage pattern of the public under normal and unrestricted freeway conditions. The usage patterns usually change after the establishment of the work zone impacting analyses of work zone traffic congestion. Once a work zone is set up on a freeway, traffic demand for that segment of the freeway reduces in reaction to increased travel times and availability of alternate routes. Therefore, the traffic flow approaching a work zone can be expressed as

$$f_t(t) = \alpha_D(t)\alpha_S f'_t(t) \tag{4}$$

where f'_t = average traffic demand on the highway prior to the establishment of the work zone; α_S = seasonal demand factor; and $\alpha_D(t)$ $(0 < \alpha_D \leqslant 1)$ is the demand reduction factor (or diversion factor). The value of the latter factor depends on the number of motorists: (1) choosing alternate routes; (2) changing their schedules to avoid the work zone; (3) canceling their trips because of the work zone; and (4) changing transportation mode, for example, opting to use public transportation. Diversion of traffic through alternate routes may be signed or advised by an advanced traveler information system (ATIS). Or, it may occur naturally where motorists familiar with the highway corridor select alternate paths to their destinations. The magnitude of $\alpha_D(t)$ is determined from traffic demand studies carried out for similar freeway work zone scenarios. Alternatively, or in addition to traffic demand studies, the demand reduction factor can be determined from a traffic network analysis of the freeway corridor that includes the work zone and alternate origin-destination routes. This factor can change from one time period to another as the congestion caused by the work zone changes. This is because long queues and delays dissuade motorists from continuing on the freeway and force them to seek alternate routes.

Neural Network Model for Estimating Work Zone Capacity

Factors Affecting Work Zone Capacity and Included in Model

Accurate estimates of work zone capacity are critical for reliable and accurate computation of queue delays and lengths at work zones. Work zone capacities and, in general, freeway capacities depend on the prevailing roadway, traffic, and control conditions. As these conditions within a work zone are significantly different from those in an unrestricted segment of the freeway, work zone capacities have to be estimated separately for each work zone scenario by taking into consideration the unique characteristics of the work zone TMP that impact capacity. For example, work zone layouts may contain uncommon geometries such as lane drop-offs and sharp horizontal alignment changes that cannot support high speeds, thus reducing the freeway capacity.

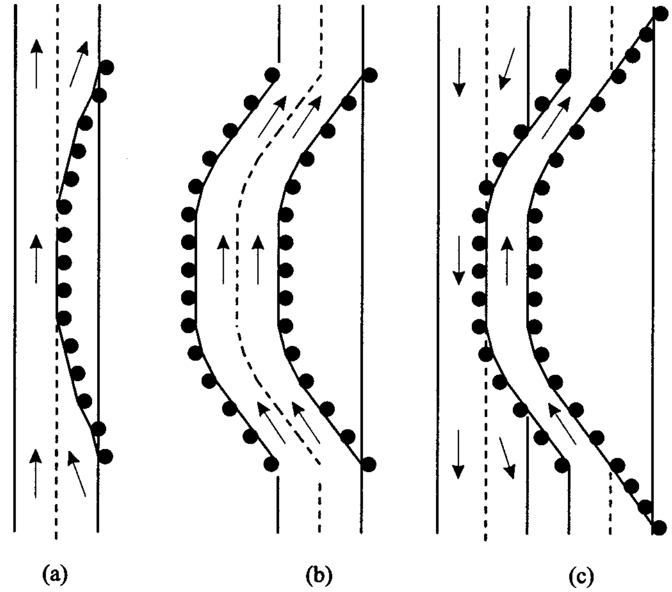

Fig. 2. Common work zone layouts: (a) lane merging; (b) lane shifting; (c) crossover

The primary factors impacting the work zone capacity and considered in this research are: number of lanes (x_1); number of open lanes (x_2); work zone layout (x_3); work zone length (x_4); lane width (x_5); percentage of trucks (x_6); grade (x_7); work zone speed (x_8); work intensity (x_9); darkness factor (x_{10}); and proximity of ramps or interchange effects (x_{11}). Theoretically, work zone capacity can be expressed as a function of these parameters:

$$c_W = f(x_i, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}) \tag{5}$$

The number of open lanes can vary from 1 to the maximum number of existing lanes in each direction. Some work zone layouts do not involve a reduction in the number of open lanes after the creation of the work zone [Fig. 2(b)]. The work zone layout parameter identifies one of the three common work zone layouts used in practice, known as lane merging, lane shifting, and crossover (shifting a diverted lane onto the right-of-way of the opposing traffic) (Fig. 2). For computational modeling, these three layouts are identified by numbers 0.1, 0.5, and 0.9, respectively. The lane width parameter is the minimum width of a traveled lane within the work zone. The standard highway lane width in the United States is 12 ft A width less than 12 ft will negatively impact the lane's capacity. The percentage of trucks parameter represents the percentage of trucks, buses, and heavy vehicles in the traffic stream. A greater percentage of trucks tends to reduce mean speeds through the work zone and consequently reduce freeway capacity. Freeway grade impacts the mean speed of traffic through the work zone, especially when the percentage of trucks is large.

As part of a TMP, a lower speed limit may be imposed and enforced through the work zone to enhance safety. A lower work zone speed, however, decreases the capacity of the work zone. The work intensity parameter describes the intensity of work activity carried out in the work zone. This is a readily nonquantifiable parameter. In this research, the work intensity is broadly categorized into low, medium, or high, depending on the size and number of the equipment and labor at the site, the noise and dust created, and the proximity of work to the traveled lanes. For example, a pavement marking operation is a low intensity work, pavement resurfacing is a medium-intensity work, and pavement

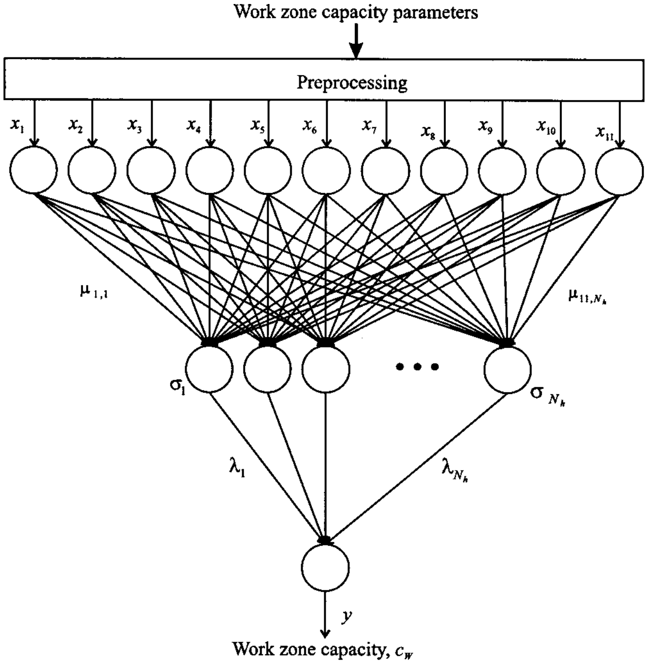

Fig. 3. RBFNN model for work zone capacity estimation

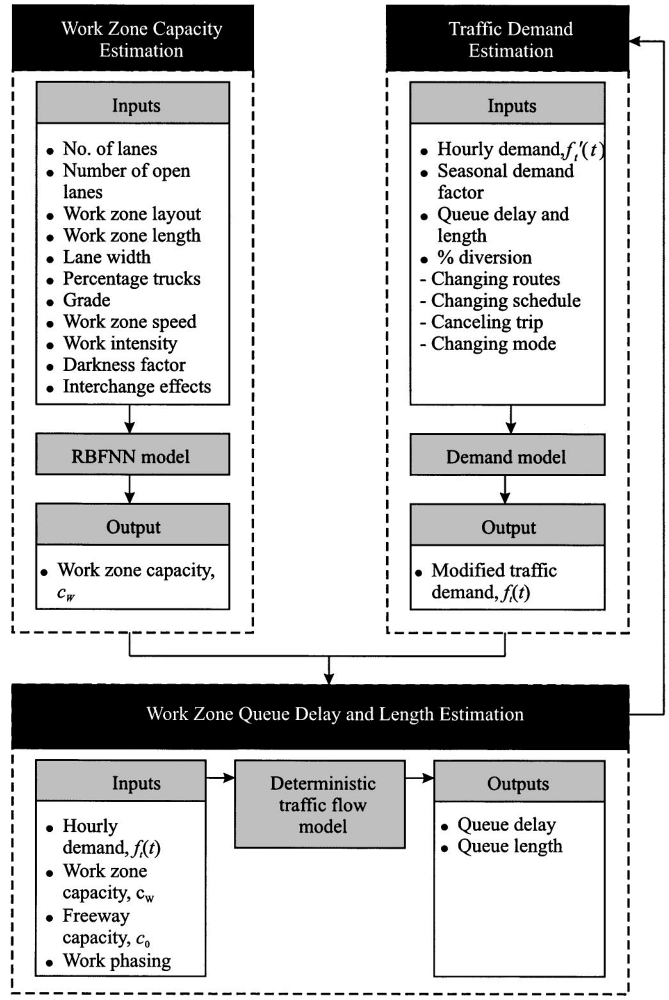

Fig. 4. Interrelation of work zone capacity and delay estimation models

rehabilitation/lane addition is a high-intensity work. For computational modeling, these three work intensity categories are identified by 0.1, 0.5, and 0.9, respectively. Low-visibility and darkness reduces capacity, as motorists become more cautious under such conditions. The darkness factor can vary from greater than 0 to 1, where 1 indicates adequate illumination that does not reduce capacity.

The interchange effects (proximity of ramps) parameter indicates the presence of an on- or off-ramp within 1,500 ft upstream of the work zone taper, or 500 ft downstream of the work zone. The proximity of ramps produces turbulence in traffic flow, reducing the number of vehicles moving through the work zone. This parameter is modeled as a binary two-valued parameter with values of 0 (no ramps) or 1 (ramps exist in proximity to the work zone).

The Case for Neural Networks

There is no mathematical function for the work zone capacity function represented by Eq. (5). In other words, the work zone capacity problem is too complicated to be amenable to classical mathematical solutions. The widely used *Highway Capacity Manual* (HCM) (TRB 2000) provides scant information on the work zone capacity based on empirical data measurements. It provides a base capacity value for ideal unrestricted highway segments. This value can be modified to take into account certain deviations from the ideal conditions by applying reduction factors. Conditions within work zones are far from ideal, and the values and reduction factors given for the ideal highway segments are generally not applicable to work zone analysis. The HCM provides a base capacity of 1,600 vehicles per hour per lane (vphpl) for short-term work zones of any layout. Guidelines are also given on how to modify the base value to take into consideration work intensity, percentage trucks, proximity of ramps, and lane widths. Other factors considered in this research and described in connection with Eq. (5), such as work zone layout, are not considered. Furthermore, it is important to consider the interaction of various factors on the work zone capacity. For example, darkness has a significant impact when the work is of high intensity.

Artificial neural networks have been used to solve complicated pattern recognition and estimation problems not amenable to conventional mathematical modeling (Adeli and Hung 1995; Adeli and Park 1998; Adeli and Karim 2001). Most civil engineering applications of the neural networks are based on the simple backpropagation (BP) algorithm (Adeli 2001). But the BP algorithm has shortcomings, including very slow convergence rate and problem-dependent trial-and-error selection of the learning and momentum ratios. In the next section, a radial-basis function neural network (RBFNN) is developed for estimating the freeway work zone capacity.

The RBFNN has a simple topology consisting of an input layer, a hidden layer of nodes with radial basis transfer functions, and an output layer of nodes with linear transfer functions. Using the training data, the RBFNN creates clusters for similar patterns. Each cluster has a center (represented by a hidden layer node). Similarity of any new pattern to the training patterns is measured by its proximity to the centers of the clusters. As such, the RBFNN is a regularization or generalization network (Moody and Darken 1989; Poggio and Girosi 1990). It is most suitable for estimation problems where limited data is available and overfitting needs to be avoided. The danger of overfitting is reduced by

Table 1. Training Data for RBFNN Model for Estimating Work Zone Capacity

Number of lanes	Number of open lanes	Layout[a]	Length (mi)	Lane width (ft)	Percentage trucks	Grade (%)	Work zone speed (mi/h)	Work intensity	Darkness factor	Interchange effects	Work zone capacity (vph)
2	1	M	1	10.0	5	0	45	Low	1.00	No	1,450
3	1	M	2	11.0	5	1	45	Low	1.00	No	1,430
2	2	S	8	11.0	10	0	55	Low	0.95	No	2,900
2	2	S	1	11.0	3	1	55	Medium	1.00	Yes	2,850
3	1	C	5	11.0	8	5	50	High	1.00	No	1,350
3	2	M	2	11.0	15	0	40	Low	0.90	Yes	2,750
3	3	S	10	12.0	5	0	40	Medium	1.00	No	4,650
2	1	M	1	12.0	25	2	45	Low	1.00	Yes	1,300
2	1	C	2	12.0	15	2	35	Medium	1.00	Yes	1,250
3	2	C	5	12.0	10	0	45	High	0.90	No	2,920
4	3	M	1	10.0	5	5	35	High	1.00	Yes	4,000
3	2	M	15	10.0	7	0	40	High	1.00	No	2,750
2	2	S	20	11.0	3	1	50	Medium	0.95	No	2,950
2	2	S	3	12.0	10	0	55	Low	1.00	Yes	3,000
2	1	M	5	11.5	0	5	40	Medium	1.00	Yes	1,450
2	1	C	2	10.0	10	0	40	Low	0.95	No	1,400
3	2	M	4	11.5	15	1	45	Medium	1.00	No	2,980
3	2	M	2	12.0	25	1	45	High	1.00	No	2,650
2	1	C	2	10.0	5	0	35	Medium	0.90	Yes	1,250
2	1	M	2	10.5	8	0	45	Medium	1.00	No	1,375
2	1	M	1	12.0	3	2	55	High	1.00	No	1,550
3	2	C	5	12.0	10	1	45	Low	1.00	Yes	2,950
4	3	M	5	11.0	15	1	45	Low	1.00	Yes	4,100
3	1	M	1	10.5	20	0	35	Medium	1.00	No	1,310
2	2	S	5	10.0	10	0	40	Low	0.90	Yes	2,700
3	2	M	4	10.0	10	0	50	Medium	1.00	No	2,750
2	1	M	1	11.0	5	2	50	High	0.85	No	1,300
2	2	S	2	11.0	5	0	45	High	1.00	No	2,980
2	1	C	10	12.0	20	0	55	Medium	1.00	Yes	1,400
2	2	S	5	11.5	5	0	45	Medium	0.95	Yes	3,000
2	1	M	2	11.5	10	1	40	Low	1.00	No	1,450
2	1	M	1	10.0	5	0	35	Medium	1.00	Yes	1,300
2	2	S	2	11.0	5	5	50	Low	1.00	No	2,950
2	1	C	1	10.5	3	2	45	Low	0.90	Yes	1,420
3	1	M	1	11.0	20	0	45	Medium	1.00	No	1,380
3	2	M	1	12.0	25	1	45	Medium	1.00	No	2,750
2	2	C	5	11.0	5	1	35	Low	0.95	Yes	4,400
3	2	C	3	11.0	0	5	35	High	1.00	Yes	2,850
2	1	M	1	10.0	5	3	40	Low	1.00	No	1,350
3	1	M	1	11.5	7	0	55	Low	1.00	Yes	1,450

[a]M = lane merging; S = lane shifting; C = lane crossover.

the local nature of the transfer functions that allow only a fraction of the nodes to participate in the mapping of a given pattern. When data are limited, the effect of noise becomes significant and some patterns may not be sufficiently represented in the training. Generalization in the vicinity of cluster centers is maintained by the graded nature of the transfer functions. The generalization properties of RBFNNs are discussed in detail by Poggio and Girosi (1990) and Adeli and Wu (1998). In contrast, the multilayer feed-forward neural network trained by the BP algorithm has a large number of global transfer functions, making it susceptible to the overfitting problem when training data are limited and noisy.

Another advantage of the RBFNN over the multilayer feed-forward neural network and BP algorithm is its rapid training. Information in an RBFNN is locally distributed. As such, only a few weights have to be modified in each iteration during the training process. Because of these reasons, the RBFNN is found to be suitable for learning the work zone capacity function for which only limited data is available.

Radial-Basis Function Neural Network for Estimating Work Zone Capacity

The architecture of the proposed RBFNN for estimating the work zone capacity is shown schematically in Fig. 3. It has an input layer with eleven nodes representing the eleven parameters included in the work zone capacity function defined by Eq. (5), a hidden layer with N_h nodes with radial-basis transfer functions, and an output layer with one node representing work zone capacity.

Some of the variables in Eq. (5), such as the work intensity, are in linguistic terms. Such linguistic variables are preprocessed first by converting them to numerical values normalized between zero and one. The numerical parameters are normalized between 0 and 1 as well. The normalization is done so that no single factor dominates the training process. The number of nodes in the hidden layer, N_h, is equal to the number of cluster centers used to characterize the training data. It is chosen as a fraction of the total

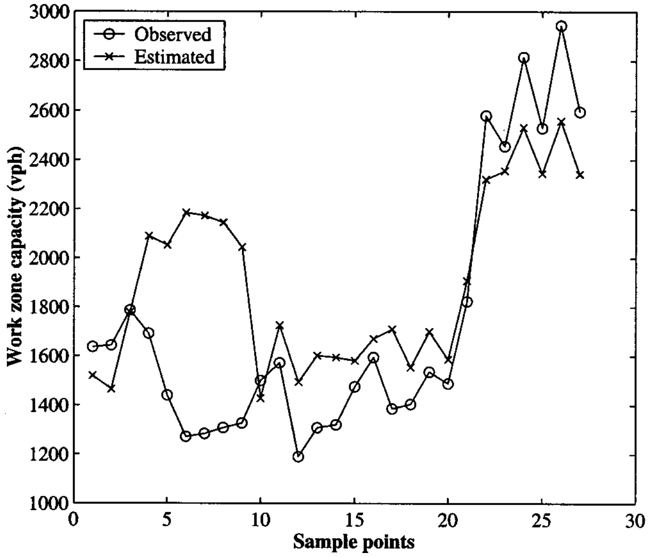

Fig. 5. Comparison of observed and RBFNN estimated capacity values

number of training instances. This choice is based on numerical experimentation for the problem at hand to determine which number adequately covers the input space and produces the best mapping. A number within the range of 10–30% of the number of training instances is found to produce satisfactory results (Adeli and Karim 2000). The cluster centers are represented by vectors $\boldsymbol{\mu}_j$ $(1 \leqslant j \leqslant N_h)$ obtained using the fuzzy c-means algorithm described in Adeli and Karim (2000).

The weight of the link connecting the input node i to the hidden node j is set equal to μ_{ij} corresponding to the ith component of the vector $\boldsymbol{\mu}_j$. The output of a hidden node j is determined by the following Gaussian transfer function:

$$\phi_j = \exp\left(-\frac{\|\mathbf{x}-\boldsymbol{\mu}_j\|^2}{2\sigma_j^2}\right) \tag{6}$$

where $\mathbf{x}$ = vector of input variables; and the factor σ_j controls the spread or range of influence of the Gaussian function centered at $\boldsymbol{\mu}_j$. In this work σ_j is calculated as (Adeli and Karim 2000)

$$\sigma_j = \frac{1}{3N_h}\sum_{i=1}^{N_h} \|\boldsymbol{\mu}_j - \boldsymbol{\mu}_i\| \quad 1 \leqslant j \leqslant N_h \tag{7}$$

Eq. (7) approximates the spread parameter σ_j as one-third of the mean distance between cluster centers. The connection from the hidden node j to the output node is assigned the weight λ_j, to be described shortly. The output y of the network is then given by

$$y = \sum_{j=1}^{N_h} \phi_j \lambda_j \tag{8}$$

The weights λ_j are calculated by minimizing the error between the network computed output y and the desired output y_d based on training examples. In other words, to train the network for λ_j's, we solve the following unconstrained optimization problem:

$$\text{Minimize} \quad E(\boldsymbol{\lambda}) = \sum_{i=1}^{N} |y^i - y_d^i| \tag{9}$$

where N = total number of training instances. The gradient descent optimization algorithm is used to solve this optimization problem. The output y of the system represents the work zone capacity, c_W.

Computation of Work Zone Queue Delays and Lengths

Fig. 4 shows schematically the interrelation of work zone capacity, traffic demand, and delay estimation models. The work zone capacity estimation model takes as input the 11 parameters affecting work zone capacity. These values are passed through the trained RBFNN model to determine the work zone capacity for the given work zone scenario. The traffic demand estimation model takes as input the hourly traffic demand prior to the establishment of the work zone, the queue delay and length, if any, and the driver behavior characteristics. The latter two factors are used to estimate the demand reduction factor. The estimates of queue delay and length are obtained from the work zone queue delay and length estimation model. This model takes as input the outputs from the traffic demand and work zone capacity estimation models. The work zone capacity and the modified traffic demand are used to estimate the work zone queue delay and length using Eqs. (3) and (2), respectively.

The model is evaluated for each time period Δt (for example, every hour). If the work zone scenario does not change during the construction, then the work zone capacity remains constant during the analysis. Otherwise, different work zone capacity values are determined for different times of the construction, reflecting the changing work zone scenario. Similarly, if the traffic flow approaching the work zone is known and the demand reduction factor is one, the traffic demand model is evaluated once. Otherwise, it is evaluated at every time step.

Training and Testing the Network

The RBFNN model for work zone capacity estimation is trained using 40 examples of work zone capacity (Table 1). These examples are created from the work zone capacity table provided by ODOT, the guidelines presented in the *Highway Capacity Manual* (TRB 2000) for different highway and work zone capacities, and the experience and judgment of the writers. The ODOT lookup tables provide work zone capacity values for lane closure configurations of 3 to 2, 3 to 1, and 2 to 1, lane widths of 10, 11, and 12 ft, truck percentages from 0 to 25, work zone lengths from 0 to greater than 8 miles, and roadway upgrades from 0 to 6%. Guidelines from the HCM, in the form of recommended ranges of adjustments, are used in conjunction with the judgement of the writers to adjust the capacity values for work zone layout, work zone speed, work intensity, and proximity of ramps. The training examples represent as many possible combinations of work zone scenarios to ensure that the boundaries of the input space are adequately covered for effective generalization by the RBFNN model. Training of the network took less than 10 s on a Pentium 4 PC with a root mean square error of 165. The purpose of the training was to achieve a generalized mapping rather than a perfect fit to the training data. A perfect fit would be of little practical value, considering the noisy nature of the training data and the complexity of the problem, which depends on many factors in addition to the eleven primary factors modeled in this research.

The network is tested using 27 work zone capacity values observed in the field and reported in the literature (denoted by scenarios 1–27 in Fig. 5). Nine samples are taken from Dixon and Hummer (1995) (scenarios 1–9 in Fig. 5). These samples are for two-lane (in each direction) rural freeways in North Carolina with one lane closed and one lane open to traffic. The North Carolina data contain 8 of the 11 parameters needed for the work zone

Table 2. Description of Work Zone Scenarios and Their Capacities Estimated by RBFNN Model

Scenario	Number of open lanes	Layout[a]	Lane width (ft)	Length (mi)	Work zone speed (mi/h)	Interchange effects	Work zone capacity (vph)
1	2	M	11.0	1	45	No	2,785
2	2	M	11.0	5	45	No	2,705
3	2	C	12.0	5	55	No	2,952
4	2	C	10.5	10	45	No	2,625
5	2	C	11.0	10	45	Yes	2,603
6	1	C	12.0	5	50	No	1,478

[a]M = lane merging; C = lane crossover.

capacity estimation model. Values of zero are used for the unavailable parameters. Twelve samples are taken from Jiang (1999) for work zones on four-lane freeways in Indiana (scenarios 10–21 in Fig. 5). These samples are also for work zones with one open and one closed lane. They contain values for only 7 of the 11 parameters. Six samples are taken from Kim et al. (2001) for work zones on freeways in Maryland (scenarios 22–27) in Fig. 5). These samples are for work zones with two open lanes. The Maryland data also contain 7 of the 11 parameters used in the new work zone capacity estimation model.

The RBFNN is used to estimate the work zone capacity for these real work zone scenarios from a variety of sites with different traffic and geometric characteristics. The work zone capacity values estimated by the RBFNN model are compared with the observed values in Fig. 5. The error is mostly in the range of 0.4–11% while for 10 samples the errors range from 20 to 71% (samples 4–9, 12–14, and 25). One reason for the large errors in the estimated work zone capacity values is the large percentage of trucks in these samples' scenarios (19–27% for samples 4–9, 22–32% for samples 12–14, and 28% for sample 25). The percentage of trucks parameter significantly impacts work zone capacity by reducing mean speeds through work zones. The impact is compounded when the work zone is on an upgrade, a parameter not reported in the sample data used to train the neural network model. Furthermore, the maximum value for the percentage of trucks in the training data set is 25, thus forcing the RBFNN model to extrapolate for scenarios with higher values. It should also be noted that the data used for testing the RBFNN model suffered from several problems, including missing values for several key parameters affecting work zone capacity and differing and/or unknown procedures for data collection and analysis. Nevertheless, considering the very limited amount of training data available to the writers, the model yields reasonably accurate results for most scenarios.

The RBFNN model presented is general and is expected to perform much better when tested with a larger and more reliable data set. This observation is based on the fact that the training root mean square error is 165 vehicles per hour, which is acceptable for most practical purposes.

Table 3. Queue Delay Results for Example 1

Hour of day	Traffic demand (vph)	Number of vehicles in queue: Work zone scenarios (defined in Table 2)					
		1	2	3	4	5	6
0	682	0	0	0	0	0	0
1	431	0	0	0	0	0	0
2	304	0	0	0	0	0	0
3	323	0	0	0	0	0	0
4	312	0	0	0	0	0	0
5	580	0	0	0	0	0	0
6	1,934	0	0	0	0	0	456
7	2,986	201	281	34	361	383	1,964
8	2,666	82	242	0	402	446	3,152
9	3,067	364	604	115	844	910	4,741
10	2,681	260	580	0	900	988	5,944
11	3,035	510	910	83	1,310	1,420	7,501
12	2,887	612	1,092	18	1,572	1,704	8,910
13	2,761	588	1,148	0	1,708	1,862	10,193
14	3,133	0	0	0	0	0	7,926
15	3,503	0	0	0	0	0	6,029
16	3,586	0	0	0	0	0	4,215
17	4,027	0	0	0	0	0	2,842
18	2,609	0	0	0	0	0	51
19	1,895	0	0	0	0	0	0
20	1,591	0	0	0	0	0	0
21	1,492	0	0	0	0	0	0
22	1,423	0	0	0	0	0	0
23	833	0	0	0	0	0	0
Maximum queue length (mi)		1	1.9	0.19	2.8	3.1	17.0

Table 4. Queue Delay Results for Example 2

Hour of day	Demand reduction factor	Number of vehicles in queue (with demand reduction)	Number of vehicles in queue (from Table 3, Scenario 1)
0	1.00	0	0
1	1.00	0	0
2	1.00	0	0
3	1.00	0	0
4	1.00	0	0
5	0.99	0	0
6	0.98	0	0
7	0.95	52	201
8	0.95	0	82
9	0.97	190	364
10	0.97	6	260
11	0.95	104	510
12	0.94	33	612
13	0.96	0	588
14	1.00	0	0
15	1.00	0	0
16	1.00	0	0
17	1.00	0	0
18	1.00	0	0
19	1.00	0	0
20	1.00	0	0
21	1.00	0	0
22	1.00	0	0
23	1.00	0	0

Examples

Example 1

This example demonstrates the use of the new work zone capacity and delay estimation model as an intelligent decision support system for the creation of a work zone TMP. It also highlights the significance of accurately estimating the work zone capacity for reliable estimation of queue delay and length. A work zone needs to be established on a six-lane freeway (three lanes in each direction) for lane resurfacing. The work will start at 6 a.m. each day and terminate 8 h later. The capacity of the unrestricted freeway (in one direction) is 5,400 vph.

Six work zone scenarios are evaluated for this work, as described in Table 2. Each scenario represents a different work zone geometry and management option. It is desired that at least two lanes be kept open through the work zone. The impact of work zone layout, lane width, length, speed, and proximity of ramps are investigated. A work zone scenario with one open lane (scenario 6) is also considered for situations where it becomes unavoidable. The RBFNN model for estimating work zone capacity is used to determine capacity values for these scenarios. The results, given in Table 2, clearly show the significant dependence of work zone capacity on parameters such as length, lane width, and proximity of ramps. By increasing the work zone length from 1 to 5 mi, the work zone capacity is reduced from 2,785 to 2,705 vph (scenarios 1 and 2). The width of lanes has a more significant effect, as seen from the estimated capacity values for scenarios 3, 4, and 5. The proximity of a ramp reduces capacity by only 22 vph (scenarios 4 and 5). Comparing scenarios 1 and 2 with scenarios 3, 4, and 5, it is seen that the lane merging layout produces slightly better capacity values as compared with the crossover layout.

Table 3 gives the number of vehicles in queue at each hour (or queue delay in vehicle-hours) for the six work zone scenarios. The hourly traffic flow approaching the work zone (traffic demand) is given in column 2. In this example, it is assumed that the traffic demand reduction factor is 1. The maximum vehicles in queue produced by the six scenarios varies from 115 (scenario 3) to 10,193 (scenario 6). Assuming the jam density (k) is 200 vehicles per mile per lane and the queue is evenly distribution among the three lanes, the maximum queue lengths produced by these scenarios vary from 0.19 to 17 mi. The maximum determined for scenario 6 may not be attained in practice, because motorists would react to the delay and change their behavior. In any case, scenario 6 should not be adopted except for emergency situations. Work zone scenario 3 provides the best solution for this example. This scenario satisfies ODOT's requirement that queue lengths should be less than 0.75 mi long at all times (ODOT 2000).

The significance of accurate work zone capacity estimation is evident from Table 3. Comparing scenarios 3 and 4, a slight increase in work zone capacity caused by modifying the work zone characteristic has drastically reduced the queue delay. Scenario 4 causes a daily queue delay of 7,097 vehicle-hours, as compared to only 250 vehicle-hours for scenario 3.

Example 2

This example illustrates the use of the demand reduction factor and its impact on the computation of work zone queue delay and length. Hourly traffic demand is often measured on or estimated for unrestricted freeways. However, when a work zone is established, the traffic flow approaching the work zone often reduces, as motorists change behavior in reaction to delays or delay information. Data for the six-lane freeway presented in Example 1 is used in this example. Work zone scenario 1 is analyzed. The traffic demand on the freeway and the demand reduction factors are given in Table 4. Up to 6% of motorists change their behavior and reduce the flow approaching the work zone. The results of the capacity and delay estimation model are given in Table 4. It is seen that the number of vehicles in queue decreases sharply as compared with the case when no reduction in demand is considered in the analysis. This example highlights the importance of warning motorists in advance and providing them with alternate routes in reduction of queue delays and lengths. It also shows that using traffic demand on unrestricted freeways for the computation of work zone queue delays and lengths will overestimate these values.

Table 5. Queue Delay Results for Example 3

Hour of day	Traffic demand (vph)	Number of vehicles in queue		
		Phasing 1	Phasing 2	Phasing 3
0	180	0	0	0
1	50	0	0	0
2	117	0	0	0
3	420	0	0	0
4	833	0	0	0
5	1,145	0	0	0
6	2,161	0	580	0
7	821	0	0	0
8	1,020	0	0	0
9	930	0	0	0
10	910	0	0	0
11	1,320	0	0	0
12	1,620	39	0	0
13	1,728	186	0	0
14	2,154	759	0	0
15	2,420	1,509	0	0
16	2,021	2,038	0	0
17	1,460	1,917	0	0
18	850	0	0	0
19	700	0	0	0
20	400	0	0	0
21	280	0	0	0
22	240	0	0	0
23	210	0	0	0

Example 3

This example illustrates the impact of work scheduling on work zone delays and queue lengths. A work zone on a four-lane (two lanes in each direction) freeway is analyzed. The work zone has a lane merging layout with one lane open, having a width of 11 ft Ten percent of the traffic stream is composed of heavy vehicles. The speed limit through the work zone is 45 mi/h. The work is of medium intensity with a duration of 6 h, and no ramps exist in the proximity of the work zone. Using the RBFNN model for work zone capacity estimation, the capacity of this work zone scenario is found to be 1,581 vph. The unrestricted freeway capacity is 3,800 vph. The hourly traffic flow approaching the work zone (traffic demand) is given in Table 5. Three work zone starting times (or phasing) are analyzed. In phasing 1, work starts at noon and ends 6 h later. In phasing 2, work starts at 4 a.m., while in phasing 3 work starts at midnight. The results computed using the work zone capacity and delay estimation model are given in Table

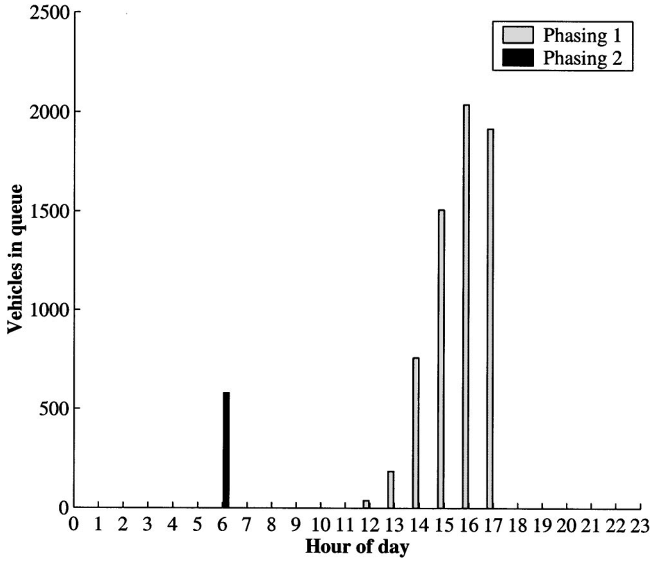

Fig. 6. Number of vehicles in queue for different work phasings (Example 3)

5 and shown in Fig. 6. It is seen that, by scheduling work when traffic demand is low, one can reduce or even eliminate work zone queue delays and lengths..

Conclusion

In this paper, a new model for work zone capacity and delay estimation is presented. The model considers 11 parameters in the estimation of work zone capacity: number of lanes, number of open lanes, layout, length, lane width, percentage trucks, grade, speed, work intensity, darkness factor, and proximity of ramps. A RBFNN model is developed to map a work zone scenario to its capacity. The model also considers reduction in traffic flow approaching the work zone in the computation of queue delays and lengths. Accurate estimation of work zone capacity and demand are essential for the accurate and reliable determination of work zone queue delays and lengths. Three examples are presented to illustrate the use of the new model and to highlight the significance of capacity and demand values in the analysis of work zones.

Acknowledgments

This manuscript is based on a research project sponsored by the Ohio Department of Transportation and the Federal Highway Administration.

References

Adeli, H. (2001). ''Neural networks in civil engineering: 1989-2000.'' *Comput. Aided Civ. Infrastruct. Eng.*, 16(2), 126–142.

Adeli, H., and Hung, S. L. (1995). *Machine learning—neural networks, genetic algorithms, and fuzzy systems*, Wiley, New York.

Adeli, H., and Karim, A. (2000). ''A fuzzy-wavelet RBFNN model for freeway incident detection.'' *J. Transp. Eng.*, 126(6), 464–471.

Adeli, H., and Karim, A. (2001). *Construction scheduling, cost optimization, and management: a new model based on neurocomputing and object technologies*, E&FN Spon, London.

Adeli, H., and Park, H. S. (1998). *Neurocomputing for design automation*, CRC Press, Boca Raton, Fla.

Adeli, H., and Wu, M. (1998). ''Regularization neural network for construction cost estimation.'' *J. Constr. Eng. Manage.*, 124(1), 18–24.

Cassidy, M. J., and Bertini, R. L. (1999). ''Some traffic features at freeway bottlenecks.'' *Trans. Res., Part B: Methodol.*, 33, 25–42.

Cassidy, M. J., and Mauch, M. (2001). ''An observed traffic pattern in long freeway queues.'' *Trans. Res., Part A: Policy Pract.*, 35, 143–156.

Chien, S., and Schonfeld, P. (2001). ''Optimal work zone lengths for four-lane highways.'' *J. Transp. Eng.*, 127(2), 124–131.

Cottrell, W. D. (2001). ''Empirical freeway queuing duration model.'' *J. Transp. Eng.*, 127(1), 13–20.

Dixon, K. K., and Hummer, J. E. (1995). *Capacity and delay in major freeway construction*, Center for Transportation Engineering Studies, North Carolina State University, Raleigh, N.C.

Dixon, K. K., Hummer, J. E., and Lorscheider, A. R. (1996). ''Capacity for North Carolina freeway work zones.'' *Transportation Research Record 1529*, Transportation Research Board, Washington, D.C., 27–34.

Federal Highway Administration (FHWA). (1998). ''Transportation Equity Act for the 21st century.'' ⟨http://www.fhwa.dot.gov/tea21⟩.

Federal Highway Administration (FHWA). (2000). ''Meeting the customer's needs for mobility and safety during construction and maintenance operations.'' ⟨http://www.fhwa.dot.gov/reports/bestprac.pdf⟩.

Islam, M. N., and Seneviratne, P. N. (1993). ''Work-zone traffic management with transportation planning software.'' *Can. J. Civ. Eng.*, 20, 471–479.

Jiang, Y. (1999). ''Traffic capacity, speed, and queue-discharge rate of Indiana's four-lane freeway work zones.'' *Transportation Research Record 1657*, Transportation Research Board, Washington, D.C., 10–17.

Kim, T., Lovell, D. J., and Paracha, J. (2001). ''A new methodology to estimate capacity for freeway work zones.'' *Proc., 2001 Transportation Research Board Annual Meeting* ⟨http://wzsafety.tamu.edu/docs/00675.pdf⟩.

Krammes, R. A., Dudek, C. L., and Memmott, J. L. (1987). ''Computer model for evaluating and scheduling freeway work-zone lane closures.'' *Transportation Research Record 1148*, Transportation Research Board, Washington, D.C., 18–24.

Krammes, R. A., and Lopez, G. O. (1994). ''Updated capacity values for short-term freeway work zone lane closures.'' *Transportation Research Record 1442*, Transportation Research Board, Washington, D.C., 49–56.

May, A. D. (1990). *Traffic flow fundamentals*, Prentice-Hall, Englewood Cliffs, N.J.

Memmott, J. L., and Dudek, C. L. (1984). ''Queue and user cost evaluation of work zones (QUEWZ).'' *Transportation Research Record 979*, Transportation Research Board, Washington, D.C., 12–19.

Mitretek. (2000). ''QuickZone delay estimation program: user guide, beta version 0.91.'' ⟨http://www.ops.fhwa.dot.gov/wz/quickz.htm⟩.

Moody, J., and Darken, C. J. (1989). ''Fast learning in networks of locally-tuned processing units.'' *Neural Comput.*, 1, 281–294.

Ohio Department of Transportation (ODOT). (2000). ''Traffic management in work zones: interstate and other freeways.'' *Policy No. 516-003(P)*, Columbus, Ohio.

Poggio, T., and Girosi, F. (1990). ''Networks for approximation and learning.'' *Proc. IEEE*, 78, 1481–1497.

Sadegh, A., Radwan, A. E., and Rouphail, N. M. (1988). ''ARTWORK: a simulation model of urban arterial work zones.'' *Transportation Research Record 1163*, Transportation Research Board, Washington, D.C., 1–3.

Transportation Research Board (TRB). (2000). *Highway capacity manual*, Washington, D.C.

Ullman, G. L. (1996). ''Queuing and natural diversion at short-term freeway work zone lane closures.'' *Transportation Research Record 1529*, Transportation Research Board, Washington, D.C., 19–26.